DOCUMENT RESUME

ED 308 219 TM 013 567

AUTHOR Goldberg, Gail Lynn; Walker-Bartnick, Leslie
TITLE Designing a Review and Appeal Process for a Large Scale Writing Assessment Program.
PUB DATE Mar 89
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cutting Scores; Evaluators; *Formative Evaluation; *Graduation Requirements; High Schools; *High School Students; Interviews; Pilot Projects; Scores; Secondary Education; Standardized Tests; State Programs; *Testing Programs; Test Results; Test Validity; *Writing Evaluation
IDENTIFIERS *Appeals (Hearings); Large Scale Programs; *Maryland Writing Test; Retesting

ABSTRACT

A formative evaluation of a pilot review and appeal process for the Maryland Writing Test (MWT) is described. The MWT is a large-scale direct assessment of writing. A three-year pilot phase was to culminate in the implementation of an operational procedure for the review and appeal of scores impacting a pass/fail decision for examinees. The MWT requires that students respond to a pair of three-paragraph prompts, one narrative and one explanatory. Each response is scored via a modified holistic procedure on a domain-specific rubric covering five criteria--content, organization, audience, sentence formation, and conventions of edited American English. During the first 2 years of the piloting of the MWT (1987 and 1988), approximately 50 reviewers were selected by the local education agencies for training by the State Department of Education; the reviewers were trained to conduct review sessions with students and/or parents. Review sessions were designed to help parents and students understand MWT scores and explanatory rubrics (scoring guides) and to provide information to help participants decide concerning possible appeals. Pilot study data from 1987 and 1988 indicate that the MWT scores have a high degree of validity. Of the 15,955 10th graders tested in 1987, 2,392 received a total raw score of 5.0 on the MWT, and thus were eligible for review and appeal. Of the 23,331 10th and 11th graders tested in 1988, 2,951 received a total raw score of 5.0, and thus were eligible for review and appeal. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED308219

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

GAIL LYNN GOLDBERG

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

DESIGNING A REVIEW AND APPEAL PROCESS FOR A LARGE SCALE WRITING ASSESSMENT PROGRAM

Gail Lynn Goldberg

Leslie Walker-Bartnick

Maryland State Department of Education

Paper presented at the annual meeting of the American Educational Research Association, March 27-31, 1989, San Francisco, California

TM013567

BEST COPY AVAILABLE

Designing a Review and Appeal Process for a Large Scale Writing Assessment Program

Gail Lynn Goldberg and Leslie Walker-Bartnick
Maryland State Department of Education

Introduction

The purpose of this study was to describe and conduct a formative evaluation of a pilot review and appeal process for the Maryland Writing Test (MWT), a large scale direct assessment of writing. A three year pilot phase was to culminate in the implementation of an operational procedure for the review and appeal of scores impacting a pass/fail decision. This investigation was an essential component of activities intended to result in a valid and equitable means of confirming or correcting the scores given on the MWT.

Introduced in 1984, the MWT requires that students respond to a pair of three-paragraph prompts, one narrative and one explanatory prompt. Each response is scored using a modified holistic procedure on a domain-specific rubric covering five criteria--content, organization, audience, sentence formation and conventions of edited American English. Well before 1987, the anticipated date by which the MWT was to become a graduation requirement, Maryland educators began to consider a score review and appeal procedure as a necessary adjunct to this statewide writing assessment program. The advisability of a review and appeal procedure was corroborated in 1986 by the Educational Testing Service (ETS) which conducted a review of the Maryland Writing Test at the request of the Maryland State Department of Education (MSDE). ETS recommended that the department consider a method for reexamining the responses of students scoring below the passing level, and identified the need for an appeal process. To permit adequate time for various refinements to the test, including the design and implementation of a review and appeal process, the state superintendent deferred the original date when passage of the MWT was to become a graduation requirement from 1987 to 1989. Concurrently, he charged a panel comprised of staff from local educational agencies (LEAs) and MSDE with designing the details of the review and appeal process. The Program Assessment, Evaluation, and Instructional Support Branch (PAEISB), the branch responsible for the Maryland Functional Testing Program, played a central role. Their specific task was to provide data which could inform recommendations for expansion and further refinement before implementation of the process in a permanent format by 1989.

Background

The MWT is a reliable, valid measure of students' narrative and explanatory writing skills (Maryland State Department of Education, 1986; 1987; 1988; Educational Testing Service, 1986, 1988). The scoring process used requires intensive reader training and involves numerous procedures designed to assure accurate, reliable scoring of each student's responses. These procedures include scoring by at least two readers, monitoring of readers' scoring by expert readers, and the use of prescored validity packets. Nevertheless, it was considered very important to reduce the possibility that any scoring error would keep a student from passing the test. This meant that reader accuracy was of special concern for students who received a total raw score of 5.0 (just below the passing score of 5.5). This total score reflects the combined scores on responses in both domains; each of those scores is obtained by averaging the score assigned by two trained readers using a four point scale. Since adjacent agreement is acceptable for the MWT, students may receive a 5.0 because of some combination of exact agreement scores or of "split scores." The Review and Appeal Panel envisioned eligibility as extending to all students but first-time takers for whom one or more score changes would move the total score to passing. If, upon review, students suspected a possible scoring error, they would be able to appeal and receive either confirmation or correction of their scores.

In the course of their deliberations, the Review and Appeal Panel identified three key questions: 1) Should only scores of 5.0 be eligible, or should total scores of 4.5 also be eligible? 2) Should students be permitted to appeal only one domain score? 3) Should all 5.0 scores be eligible, or only those resulting from adjacent agreement scores? The hypotheses of the PAEISB study which derived from these questions were that 1) patterns of score change will be such that only 5.0 scores will be sufficiently affected to impact a pass/fail decision; 2) score changes may occur in both domains, canceling each other out with the result of the same total test score; and 3) changes will occur primarily in adjacent agreement scores. Confirmation or rejection of each of the hypotheses over the pilot years would provide the framework for the operational review and appeal procedure.

Methodology

During each of the first two years of the pilot, immediately preceeding the return of test booklets to the LEAs, approximately 50 reviewers were selected by the LEAs for training by MSDE on conducting review sessions with students and/or parents. The agreed upon purpose of the review was to explain scores based on

the MWT narrative and explanatory rubrics (scoring guides) and to provide information to help students and their parents decide whether to appeal. Upon receipt of an appeal request, a copy of the responses was obtained and sent to the scoring contractor for rescoring. Utilizing the same scoring procedures, rubrics, and guide papers, senior staff members who were expert readers rescored the appealed responses and confirmed or corrected original scores. Adjacent agreement scores were confirmed if two expert readers could identify features of different score points in the response. Difficult appeal decisions were reviewed by PAEISB writing assessment specialists. Results were reported to students and LEAs, and the data subjected to detailed analysis by PAEISB. Of particular concern were the patterns of score change by domain, by score point, and by score type (exact vs. adjacent agreement scores).

LEAs were responsible for notifying eligible students of the opportunity for review and appeal. It was up to the students or their parents to request a review session, and then to decide whether or not to appeal. This resulted in a self-selected group of students in the review and appeal data set.

Based on the new graduation requirement date, in the first year of the pilot only tenth grade students who obtained a total raw score of 5.0 on the 1987 MWT were eligible for review and appeal. In 1987, of 15,955 tenth grade test takers, 2,392 received a total raw score of 5.0. Of these 5.0's, 2097 (87.7%) resulted from exact agreement scores, and 295 (12.3%) from adjacent agreement. The most frequent score combination contributing to a total raw score of 5.0 (77.8% of total 5.0's) was a narrative score of 3 and an explanatory score of 2. The reverse combination of exact agreement scores accounted for an additional 9.7% of the total 5.0 scores. Among adjacent agreement combinations, the most common was a narrative split score of 2.5 coupled with an explanatory score of 2.5. This combination accounted for 11.1% of the total scores of 5.0 (see Table 1).

In the second year of the pilot, both tenth and eleventh graders who obtained a total score of 5.0 on the 1988 MWT were eligible for review and appeal. Of 23,331 tenth and eleventh grade test takers, 2,951 received a total raw score of 5.0. Of these 5.0's, 2699 (91.5%) resulted from exact agreement scores and 252 (8.5%) from adjacent agreement. In 1988, the most frequent score combination contributing to a total raw score of 5.0 (85.4% of total 5.0's) again was a narrative score of 3 and an explanatory score of 2. The reverse combination of exact agreement scores accounted for an additional 5.5% of the total 5.0 scores. Among adjacent agreement combinations, the most common was a narrative split score of 2.5 coupled with an explanatory score of 2.5. In 1988, this combination accounted for 7.4 % of the total scores of 5.0 (see Table 1).

Results

1987 Results

Of the 2,392 tenth grade students eligible for review and appeal in 1987, 404 (17% of those eligible) reviewed, and 298 went on, following the review session, to appeal their scores. As a result of the appeals process, there were 45 changes in original 5.0 scores, and 36 students received scores sufficient to pass the MWT. Among those students with score changes which resulted from an appeal rescoring, the total score was raised to 5.5 in 23 instances and to 6.0 in 13 instances. Among students whose scores were lowered, the resultant score was 4.5 in six instances and 4.0 in three instances. In four instances, a score change made in one domain was cancelled out by a reverse score change in the other domain (see Table 2).

Of the students who received a new, passing score on the MWT, twenty did so because of a change in their narrative score, fifteen because of a change in their explanatory score, and one because of a change in both domain scores. Examination of the number of score changes by domain reveals that the majority of instances of reverse score changes leading to retention of the original total score of 5.0 occured due to a raised narrative score combined with a lowered explanatory score (see Table 3). Although the same number of score changes occured in both domains, of all appealed responses rescored, 2% more narrative scores went up than explanatory scores, and 2% more explanatory scores went down than narrative scores (see Table 4).

Of the 298 students who appealed in 1987, 44 (15%) had total original 5.0s that were the result of adjacent agreement scores (split scores of 1.5, 2.5, or 3.5 in either domain totaling 5.0) while 254 (85%) had original scores of 5.0 that were the result of some combination of solid scores in each domain (scores of 1, 2, 3, or 4) (see Table 5). Among the 36 new passers, 24 benefited from a change in adjacent agreement scores while 12 benefited from a change in solid scores.

1987 Summary

Examination of score results from the first year of the pilot review and appeal procedure provided preliminary support for the hypotheses underlying the investigation. Hypothesis 1 states that score change patterns would suggest that only 5.0 scores will be sufficiently affected to impact a pass/fail decision. Of the 49 score changes, 44 (90%) represented a change in only one of a student's scores, and the majority of changes (63% in narrative and 74% in explanatory) were from adjacent agreement scores. This means that in most cases, students passed because of the addition of only one half of a point to their 5.0 to give them a 5.5. Only 36% of the new passers had a whole point added to their total score. Based on this data, it was

decided to maintain the review and appeal eligibility total score at 5.0, and not open it to the students with total scores of 4.5.

Hypothesis 2 states that score changes may occur in both domains and even cancel each other out. Score changes occurred in both domains in 5 instances; 4 of these were cancellations, and the final one was the result of both scores going up. Although this occurred in a minority of cases, it did occur, and coupled with the philosophy that the two domains make up the test as a whole, it was decided to maintain the policy of reviewing and appealing both domain scores. There was no evidence to suggest this policy was detrimental to students.

Hypothesis 3 stated that score changes would occur primarily in adjacent agreement scores. Although this hypothesis was confirmed with 37 (69%) of the score changes occurring to adjacent agreement scores, 17 (31%) solid scores were also changed. Therefore, in the interest of fairness to students with solid scores, it was decided not to limit review and appeal to students with adjacent agreement scores. Approximately half of the adjacent agreement scores were confirmed during rescoring, providing additional support for the validity of split scores.

1988 Results

The findings in the first year of the pilot helped to shape the review and appeal procedures for 1988 and permitted PAEISB to recommend that the same criteria for eligibility be maintained, with the addition of another grade to the eligible pool. The outcome was that of 2,951 students eligible for review and appeal in 1988, 508 (17.2% of those eligible) took the opportunity to review their scores and 374 (73.6% of those who reviewed) went on, following the review session, to appeal their scores. These percentages are remarkably consistent with those from 1987. As a result of the appeal process, there were 45 changes in original 5.0 scores, and 35 students received scores sufficient to pass the MWT. Among those students with score changes which resulted from an appeal rescoring, the total score was raised to 5.5 in 13 instances and to 6.0 in 22 instances. Among students whose total score was lowered, the resultant score was 4.5 in 6 instances, 4.0 in 3 instances, and 3.5 in one instance. As in 1987, there were four instances in which a score change made in one domain was canceled out by a reverse score change in the other domain (see Table 6).

Of the students who received a new, passing score on the MWT as a result of the appeal rescoring, twenty did so because of a change in their narrative score, thirteen because of a change in their explanatory score, and two because of a change in both domain scores. Examination of the number of score changes by domain reveals that the majority of instances of reverse score changes leading to retention of the original total score of 5.0 occurred due to a raised narrative score combined with a lowered explanatory score, although this occurred with less frequency

than in 1987 (see Table 7). Unlike 1987, when there was an equal number of score changes in both domains, 1988 witnessed almost a third more narrative score changes than explanatory score changes. Among these score changes the number of narrative changes upward continued to be greater than the number of explanatory changes upward; in both domains, however, the number of score changes downward in each domain was identical in 1988 (see Table 8).

Of the 374 students who appealed in 1988, 36 (9.6%) had total original 5.0's that were the result of adjacent agreement scores, while 338 (90.4%) had original scores of 5.0 that were the result of some combination of solid scores in each domain (see Table 9). Among the 35 new passers, 20 benefited from a change in solid scores while 15 benefited from a change in adjacent agreement scores.

1988 Summary

Results of the second year of the pilot generally supported the hypotheses, but suggested the necessity to continue to follow up on some emerging trends. Relevant to hypothesis one, 96% of the score changes occurred in only one domain. However, while in the narrative domain the majority of upward changes (75%) increased the total raw score by only one half of a point, in the explanatory domain there was an increase in upward changes made to solid scores which resulted in 22 (63%) new passers having a whole point added to their total score. If this trend continues, it may be desirable to consider allowing students with total scores of 4.5 to appeal, since addition of a full point would bring their score to a passing 5.5.

The 1988 results provided some support for hypothesis 2, since the data revealed 4 instances of an increased score in one domain being canceled by a decreased score in the other. In addition, 2 students had scores go up in both domains.

Finally, the results of the narrative domain score changes, 75% of which were in adjacent agreement scores, supported hypothesis 3. While in the explanatory domain only 39% of changes occurred in adjacent agreement scores, the small number of score changes in that domain needs to be taken into consideration. While 60% of the total score changes were in adjacent agreement scores, there were 22 students who benefited from a change in a solid score. This trend suggests the need to continue to make both exact and adjacent agreement scores subject to review and appeal, and not to limit eligibility only to students with adjacent agreement scores.

Applications and Conclusions

Applications. Data from the first two pilot years of the review and appeal process confirmed the initial hypotheses and

therefore suggested that no changes were necessary in terms of cut score for eligibility, inclusion of both domain scores and both exact and adjacent agreement scores. Nevertheless, the pilot did permit PAEISB to disseminate additional information and to recommend certain changes and new features to the program for practical and instructional purposes.

Some of the most valuable information, potentially, was provided each year in the form of a chart outlining by LEA the number of students eligible for review and appeal, the number that reviewed, that went on to appeal, that had a score change, and that had a score change that affected the original pass/fail decision. While state policy prevented PAEISB from making public any data about specific LEAs, a scrambled list provided clear evidence of a considerable difference each year in those numbers by LEA. In 1988, after systems were more familiar with the program than they had been in the first year of the pilot, the percentage of eligible students availing themselves of the opportunity to review ranged from 4.5% to 78%. The percentage of students who reviewed their scores going on to appeal those scores ranged from 0% to 100%. Recognizing that these figures may reflect circumstances beyond the capacity of school systems to change, PAEISB has urged that further consideration be given to the process by which eligible students are informed of review and appeal procedures and also the the role of various types of student advocates such as teachers, parents and church leaders, in students' decision to at least seek out the information and assistance available through the review procedure. PAEISB has also built consideration of the implications of the LEA data into the annual training by MSDE of LEA reviewers.

One major procedural change from the first to second year of the pilot involved the form and contents of information returned to students who elected to appeal their original scores. In the first year of the pilot, students received a written explanation of the assigned or revised score that was composed by the contractor's expert readers. These descriptions were essentially annotations which cited particular portions of the student's text. Rather than co-opting the role of instructional personnel in providing students with descriptive analyses of test responses, PAEISB determined the need for a different type of document, the purpose of which would be to identify the features of each appealed response, based on the MWT rubrics, that contributed to the confirmed or corrected score. To facilitate this process, a computerized system was developed to create "customized" score statements for students who appealed their MWT scores.

As the pool of trained reviewers enlarges each year, PAEISB has been sensitive to changing training needs. It is likely that the format will become fixed to permit a scoring training workshop and a separate training session in which new as well as experienced reviewers can receive information on specific documents and procedures, discuss issues and concerns, review

difficult line calls and other training models, and engage in guided practice with a human relations specialist on the dynamics of a review session.

Findings also supported the need for two readers to rescore responses based on the confirmation of more than half of the adjacent agreement scores. In these instances, two or more expert readers were able to defend different scores on the basis of features of both score points evident in the responses.

Conclusions. All writing assessment programs, but particularly those involving high stakes such as certification for graduation, must endeavor to maintain fair, consistent and accurate scoring of writing samples. The very low percentages of score changes during the 2 pilot years of this study have provided additional evidence of the high level of validity of MWT scores. While the MWT has been shown to be reliable and valid, the program supports students' opportunity to verify that their individual scores are accurate. With data accrued from the two year pilot of a review and appeals process, PAEISB has been able to confirm or recommend needed modification of procedures before final implementation. For the third and final year of the review and appeal pilot, all tenth, eleventh, and twelfth grade students receiving a total raw score of 5.0 will be eligible. No other changes were deemed to be necessary at this time, except for those involving scheduling of reviews and appeal rescorings. In deference to the needs of seniors facing a graduation decison, a priority processing system will be implemented which will give seniors whose graduation might be delayed because of the MWT the opportunity to appeal in sufficient time so that a score change upward will allow those students to receive their high school diplomas in June.

Beyond contributing to the determination of eligibility, content, and context of appeal rescoring, the study has permitted confirmation of the validity of scoring guides and training materials, thus contributing to the accuracy and continuity of scoring procedures for future administrations of the MWT. By examining appeal rescoring results in light of the score distribution in each domain on that administration, statewide, PAEISB can identify where additional training is advisable. For example, as narrative mean scores increase, over time, and score points 1 and 2 become increasingly uncommon, it is likely that the 2/3 line will become harder to maintain. Appeal results will provide evidence, in this instance and others, of how successfully trained readers are able to abide by score point lines.

The 1988 pilot has also provided data on the frequency and type of features of responses at each score point and on the features which may contribute to a tendency to score incorrectly or without exact agreement. Preliminary examination suggests, for example, that the primary basis for scores of 2 in the explanatory domain is general rather than specific information, rather than insufficient specificity, organizational lapses, or

other features. Interestingly, elaboration, which is the counterpart in terms of development in narrative, appears to be less frequently the basis for lower score points. In spite of many teachers' belief that narrative is the easier domain and that students master "storytelling" far earlier than explanation, among students appealing their MWT scores inability to establish more than a minimal sequence is often the operative factor in a low score decision. It is hoped that information of this sort will be more sought after in future years, and that it will be useful to both the instructional and assessment community, within Maryland and elsewhere.

Table 1 Breakdown of 5.0 Scores, 1987 and 1988 Maryland Writing Test (MWT)

	1987			1988		
	Number	%[a]	%[b]	Number	%[a]	%[b]
Exact Agreement						
N=2,E=3	231	11.0	9.7	162	6.0	5.5
N=3,E=2	1,861	88.7	77.8	2,519	93.3	85.4
N=4,E=1	5	.2	.2	16	.6	.5
N=1,E=4	0	--	--	2	.1	.1
Total	2,097		87.7	2,699		91.5
Adjacent Agreement						
N=2.5,E=2.5	266	90.2	11.1	219	86.9	7.4
N=3.5,E=1.5	29	9.8	1.2	29	11.5	1.0
N=1.5,E=3.5	0	--	--	4	1.6	.1
Total	295		12.3	252		8.5
Total (all 5.0 scores)	2,392			2,951		

Note. 1987 includes 10th graders only (15,955 test takers); 1988 includes 10th and 11th graders (23,331 test takers).

[a]percent of total exact or adjacent agreement
[b]percent of total 5.0s

Table 2 Changes in Total Score as a Result of Appeal, 1987 (N=298)

Direction of Change	Total Score	Number	% of Total
Up	5.5	23	7.7
	6.0	13	4.4
Total		36	12.1
Down	4.5	6	2.0
	4.0	3	1.0
Total		9	3.0
No Change		253	84.9

Table 3 Analysis of Score Change by Domain, 1987

		Narrative			
		Up	Down	No Change	Total
Explanatory	Up	1[a]	1	15[a]	17
	Down	3	0	7	10
	No Change	20[a]	2	NA	22
	Total	24	3	22	49

Note. [a]Resulted in new passers.

Table 4 Analysis of Score Changes by Domain, Frequency and (Percent), 1987

	Direction of Change Up	Down	Total
Narrative	24 (8.0)	3 (1.0)	27 (9.0)
Explanatory	17 (6.0)	10 (3.0)	27 (9.0)
Total	41 (7.0)	13 (2.0)	54 (9.0)

Note. A total of 298 responses were rescored in each domain, narrative and explanatory.

Table 5 Frequency of Score Movement by Domain, 1987

Movement	Narrative # changed	Narrative # at original	Explanatory # changed	Explanatory # at original
Solid Scores				
1→2	NA	0	NA	0
2→1	0	35	0	219
2→3	5	35	5	219
3→2	2	219	2	35
3→4	3	219	0	35
4→3	NA	0	NA	0
Total	10	254	7	254
Adjacent Scores				
1.5→1	NA	0	0	4
1.5→2	NA	0	4	4
2.5→2	0	40	8	40
2.5→3	16	40	8	40
3.5→3	1	4	NA	0
3.5→4	0	4	NA	0
Total	17	44	20	44

Table 6 Changes in Total Score as a Result of Appeal, 1988 (N=374)

Direction of Change	Total Score	Number	% of Total
Up	5.5	13	3.5
	6.0	22	5.9
Total		35	9.4
Down	4.5	6	1.6
	4.0	3	.8
	3.5	1	.3
Total		10	2.7
No Change		329	88.0

Table 7 Analysis of Score Change by Domain, 1988

		Narrative			
		Up	Down	No Change	Total
Explanatory	Up	2[a]	1	13[a]	16
	Down	3	0	4	7
	No Change	20[a]	6	NA	26
	Total	25	7	17	49

Note. [a]Resulted in new passers.

Table 8 Analysis of Score Changes by Domain, Frequency and (Percent), 1988

	Direction of Change Up	Down	Total
Narrative	25 (6.7)	7 (1.9)	32 (8.6)
Explanatory	16 (4.3)	7 (1.9)	23 (6.2)
Total	41 (5.5)	14 (1.9)	55 (7.4)

Note. A total of 374 responses were rescored in each domain, narrative and explanatory.

Table 9 Frequency of Score Movement by Domain, 1988

Movement	Narrative # changed	Narrative # at original	Explanatory # changed	Explanatory # at original
Solid Scores				
1→2	NA	0	0	2
2→1	1	23	2	313
2→3	4	23	12	313
3→2	3	313	0	23
3→4	0	313	0	23
4→3	0	2	0	0
Total	8	338	14	338
Adjacent Scores				
1.5→1	0	2	0	1
1.5→2	1	2	0	1
2.5→2	4	33	5	33
2.5→3	17	33	3	33
3.5→3	1	1	0	2
3.5→4	0	1	0	2
2.5→1	1	33		
Total	24	36	9	36

References

Educational Testing Service. (1986). Review of the Maryland functional writing test program. Princeton, N.J.: Author.

Educational Testing Service. (1988). A follow-up to the ETS review of the Maryland functional writing test program. Princeton, N.J.: Author.

Maryland State Department of Education (1986). Technical report: Maryland functional writing test-II. (Report No. W-86). Baltimore, MD.: Author.

Maryland State Department of Education (1987). Technical report: Maryland functional writing test-II. (Report No. W-87). Baltimore, MD.: Author.

Maryland State Department of Education (1988). Technical report: Maryland writing test-II. (Report No. W-88). Baltimore, MD.: Author.